Electronic Supplementary Material (ESI) for Nanoscale.
This journal is © The Royal Society of Chemistry 2024

Electronic Supplementary Information

Exploiting cyclodextrins as artificial chaperones to enhance enzyme protection through supramolecular engineering

Ali Foroutan Kalourazi, Seyed Amirabbas Nazemi, Ajmal Roshan Unniram Parambil, Ruben Muñoz-Tafalla, Paula Vidal, S. Shirin Shahangian, Victor Guallar, Manuel Ferrer and Patrick Shahgaldian

Table of content

Lipase sequence-based metagenomic bioprospecting 2
Source and production of Lip_{MRD9} 2
Structural analysis for CD binding sites 4
CD-TES characterisation 5
Lip_{MRD9} immobilisation and quantification 8
SEM micrographs 9
Layer growth kinetics 10
References 11

Lipase sequence-based metagenomic bioprospecting

The hydrolase from the *Bacillus* genus (NCBI accession number: WP_034624255.1), used here, was identified during sequence-based metagenomic bioprospecting. Briefly, we screened the marine metagenomics MarRef Database,[1] containing approximately 4.7 million protein-coding sequences. The sequences were selected by querying the input sequences using DIAMOND BLASTP, with default parameters (percent identity >60%; alignment length > 70; e-value < 1×10^{-5}) against the 215 amino acid lipase from *Bacillus pumilus* (NCBI Accession Number: WP_106066877.1; molecular mass, 23,102.65 Da; isoelectric point, 9.77), a versatile lipase with industrial potential. A total of 33 sequences were retrieved (e-values from $2{,}821 \times 10^{-105}$ to 3.24×10^{-12} when compared to the lipase from *B. pumilus*). One such sequence (GenBank accession number, WP_034624255.1), assigned to a bacterium of the genus *Bacillus*, was confirmed to encode a predicted full-length 215 amino acid-long lipase (e-value $2{,}821 \times 10^{-136}$ and 92.1% similarity compared to lipase from *B. pumilus*) with the needed catalytic residues (S111, D167, H189), and it was selected as a target for further investigation. Within the MarRef Database, the sequence WP_034624255.1 originated from a microbiome isolated from the marine sponge in the seawater in front of Seongsan-ri of Jeju Island, South Korea (ENA BioSample accession SAMN06016472; ENA BioProject accession PRJNA353573). Once identified, the 215 amino acid sequence encoding the wild-type enzyme (GenBank accession number WP_075743487; molecular mass, 23,102.65 Da; isoelectric point, 9.77) was used as a template for gene synthesis. After synthesis, a 215 amino acid sequence was obtained encoding an enzyme with a molecular mass of 21,974.60 Da and an isoelectric point of 8.0. The soluble N-terminal hexahistidine (His6)-tagged protein was produced and purified (>98% using SDS-PAGE analysis; **Fig. S1**) after binding to a Ni-NTA His-Bind resin. This enzyme is referred to as Lip_{MRD9} (Lip refers to lipase; MRD refers to the MarRef Database).

Source and production of Lip_{MRD9}

The sequence of Lip_{MRD9} was synthesised by GenScript Biotech (GenScript Biotech, The Netherlands) and codon-optimised to maximise expression in *E. coli*. Before gene synthesis, the sequence was analyzed for the presence of a signal peptide using the SignalP-5.0 tool; a cleavage Sec/SPI site between positions 34 and 35 was predicted. The gene, excluding the signal peptide, was flanked by BamHI and HindIII (stop codon) restriction sites and inserted into a pET-45b(+) expression vector with an ampicillin selection marker (GenScript Biotech, The Netherlands). This vector was subsequently introduced into *E. coli* BL21 (DE3). This plasmid, which was introduced into *E. coli* BL21 (DE3), supports the expression of N-terminal His6-fusion proteins, with the final amino acid sequence of the synthetic protein being MAHHHHHHVGTGSNDDDDKSPDPM-X (where X corresponds to the original sequence of the target enzyme without the signal peptide). The soluble N-terminal His_6-tagged proteins were produced and purified (>98% purity, as determined by SDS-PAGE analysis using a Mini PROTEAN electrophoresis system, Bio-Rad, Spain) at 4 °C after binding to a Ni-NTA His-Bind resin (Merck Life Science, Spain), as previously described,[2] and stored at -20 °C until use at a concentration of 1.5 mg/mL in 40 mM 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid (HEPES) buffer (pH 7.0). Approximately 7 mg of purified protein was obtained on average from a 1-L culture.

Lip_{MRD9}: original sequence (WP_034624255.1)

MKVMFVKKRSLQILIALALVIGSMAFIQPKEVKAAEHNPVVMVHGIGGASYNFFSIKSYLATQGWDRNQLYAIDFID
KTGNNRNNGPRLSRFVKDVLDKTGAKKVDIVAHSMGGANTLYYIKNLDGGDKIENVVTIGGANGLVSSRALPGTD
PNQKILYTSVYSSADLIVVNSLSRLIGARNVLIHGVGHIGLLTSSQVKGYIKEGLNGGGQNTN

Lip_{MRD9} for gene synthesis: sequence without signal peptide, requested for gene synthesis

MAEHNPVVMVHGIGGASYNFFSIKSYLATQGWDRNQLYAIDFIDKTGNNRNNGPRLSRFVKDVLDKTGAKKVDIV
AHSMGGANTLYYIKNLDGGDKIENVVTIGGANGLVSSRALPGTDPNQKILYTSVYSSADLIVVNSLSRLIGARNVLI
HGVGHIGLLTSSQVKGYIKEGLNGGGQNTN

Lip_{MRD9} synthetic: sequence without signal peptide, requested for gene synthesis

MAHHHHHHVGTGSNDDDDKSPDPMAEHNPVVMVHGIGGASYNFFSIKSYLATQGWDRNQLYAIDFIDKTGNNR
NNGPRLSRFVKDVLDKTGAKKVDIVAHSMGGANTLYYIKNLDGGDKIENVVTIGGANGLVSSRALPGTDPNQKILY
TSVYSSADLIVVNSLSRLIGARNVLIHGVGHIGLLTSSQVKGYIKEGLNGGGQNTN

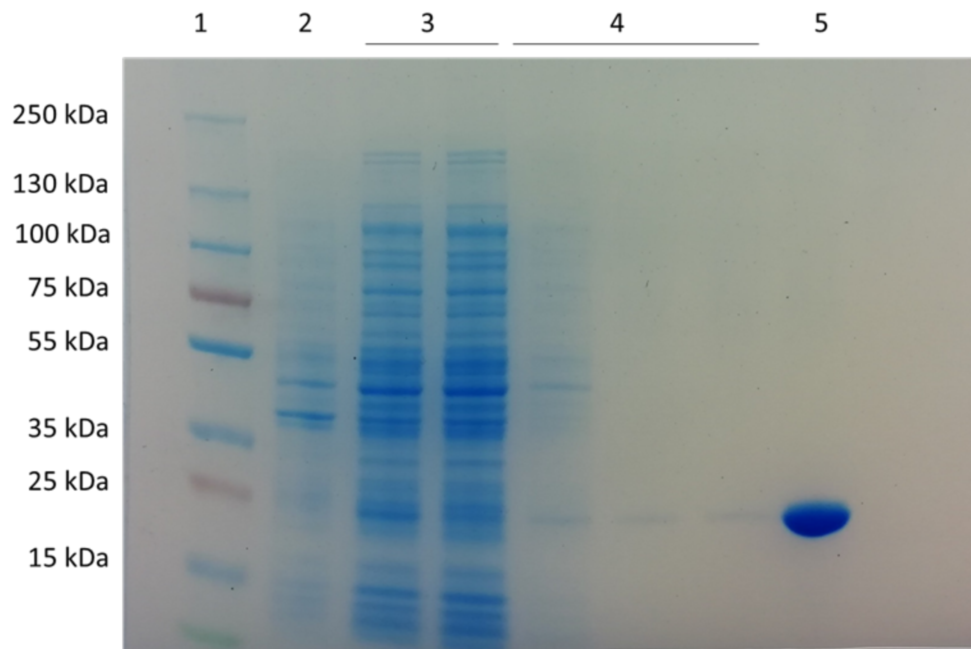

Fig. S1 0.1% SDS−15% PAGE analysis of purified Lip_{MRD9}. A total of 10 µg protein was used. Lane 1, 10-250 kDa molecular mass marker (Protein marker V (pre-stained), peqGOLD, ref. 27-2210, VWR International, Belgium); lane 2, insoluble proteins after expression; lane 3, soluble proteins after expression; lane 4, unbound proteins to Ni-NTA His-Bind resin (Merck Life Science, Spain) after different washing steps; lane 5, purified Lip_{MRD9}.

Structural analysis for CD binding sites

Table S1 Results of the structural analysis for the selected poses.

Binding Pose	Simulation Round	Binding Energy (Kcal/mol)	Ligand SASA	Residue inside the CD cavity	Hydrogen bonds number
1	2	-60.39	0.51	None	8
2	1	-59.42	0.59	Tyr	5
3	1	-51.68	0.66	Leu	8
4	1	-47.55	0.69	Asn	7
5	2	-46.25	0.57	Tyr*	8
6	2	-46.11	0.62	None	5
7	2	-43.57	0.59	Ile	5
8	2	-40.49	0.67	None	6

*Not entirely inside the cavity

CD-TES characterisation

CD-TES: ^{1}H NMR (500 MHz, DMSO-d_6): δ (ppm) 0.50-0.58 (t, 2H), 1.12-1.16 (t, 9H), 1.41-1.49 (q, 2H), 3.20-3.42 (br, 15H), 3.44-3.70 (br, 29H), 3.70-3.77 (q, 7H) 4.33-4.57 (br, 7H), 4.8-4.93 (m, 7H), 5.6-5.8 (m, 12H), 7.03 (t, 1H, NH). ^{13}C NMR (100 MHz, MeOD): δ (ppm) 8.4, 18.6, 24.2, 44.6, 59.5, 61.9, 73.7, 74.3, 74.8, 83.0, 103.6, 158.0. FT-IR (cm^{-1}): 1701 (C=O). ESI-MS: m/z calcd for $C_{52}H_{91}NO_{39}Si$ $[M+Na]^+$ 1404.4835; Found: 1404.4846.

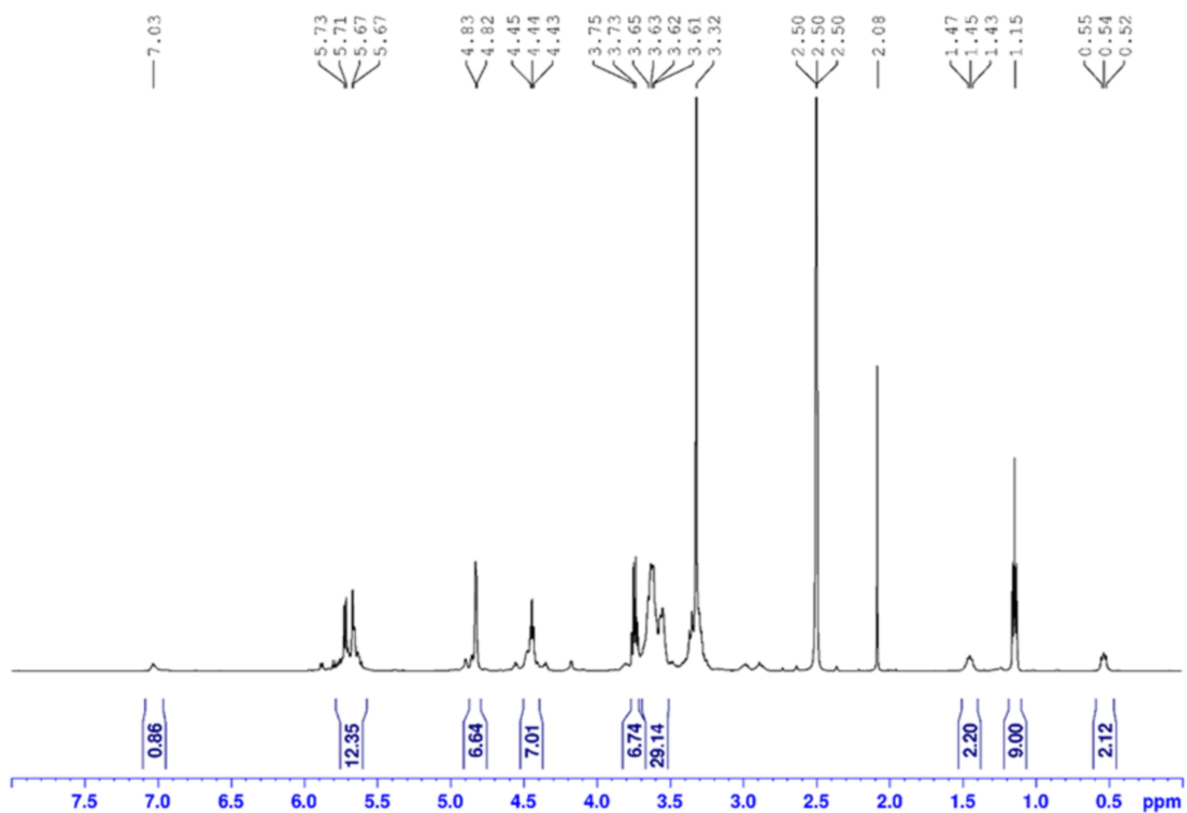

Fig. S2 ^{1}H NMR spectrum CD-TES measured in DMSO-d_6 at 500 MHz

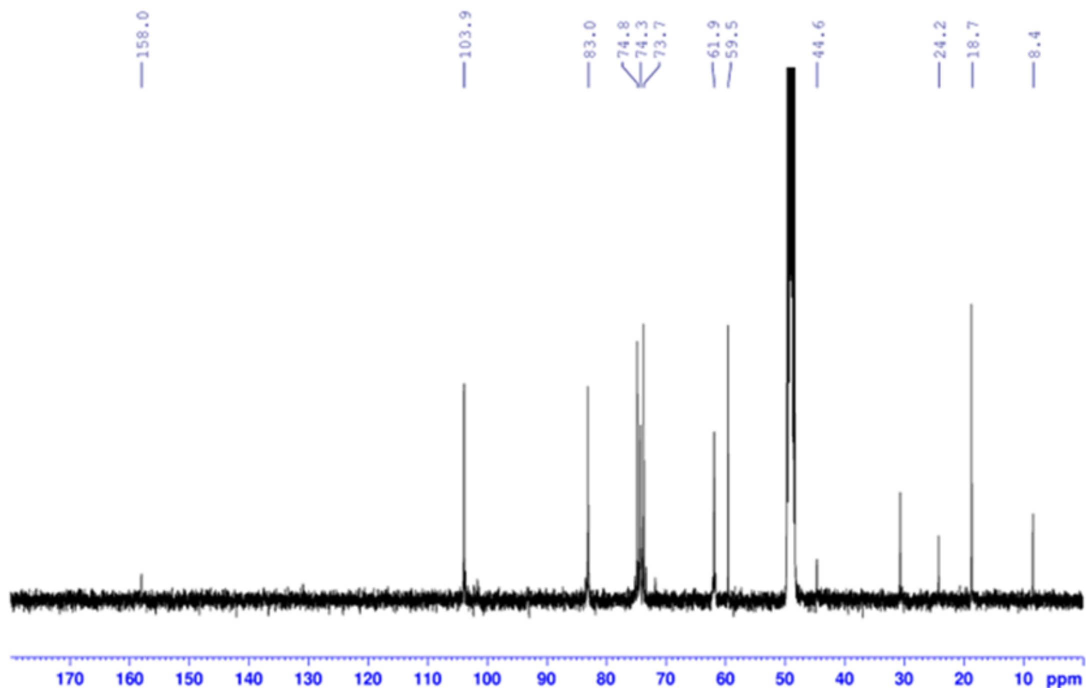

Fig. S3 ^{13}C NMR spectrum of CD-TES measured in MeOD at 100 MHz.

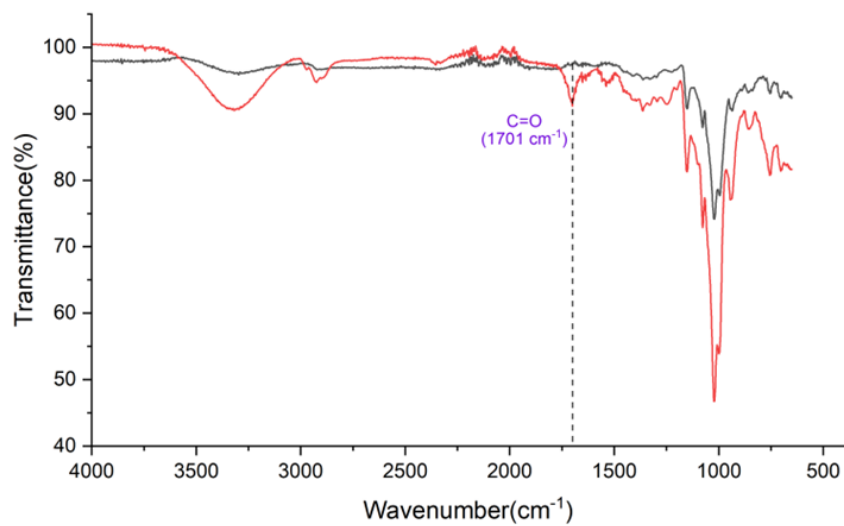

Fig. S4 FTIR spectrum of CD-TES (red) and reference CD (black).

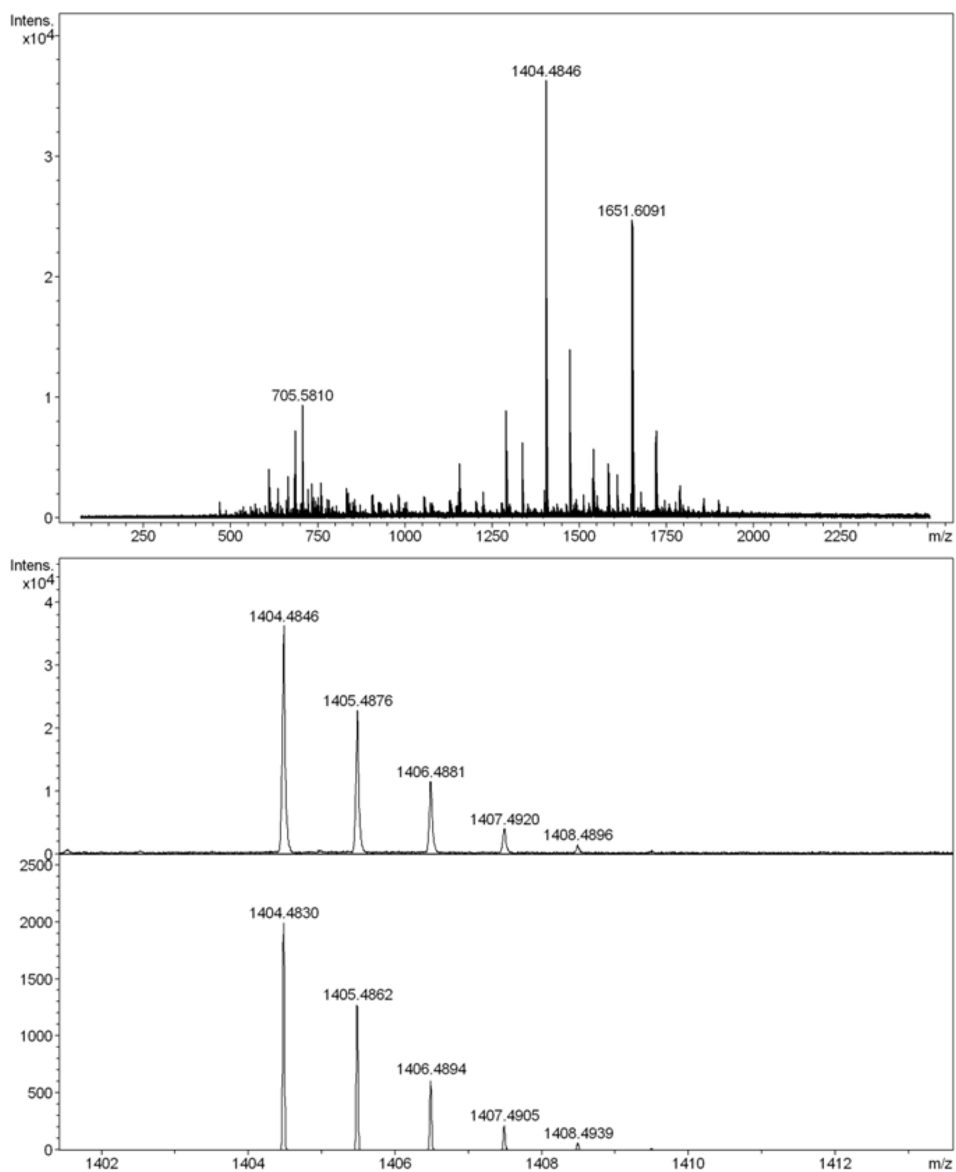

Fig. S5 ESI-MS analysis of the CD-TES.

Lip_{MRD9} immobilisation and quantification

Protein quantification assays were conducted using BCA assay kit according to the manufacturer protocol.

Table S2 Protein quantification carried out on reaction supernatants after immobilisation reaction of Lip_{MRD9}.

OD Supernatant	OD Blank	[Lip_{MRD9}] initial (µg/mL)	[Lip_{MRD9}] supernatant (µg/mL)	[Lip_{MRD9}] immobilised (µg/mL)	Average [Lip_{MRD9}] immobilised (µg/mL)	Immobilisation yield (%)	Average [Lip_{MRD9}] immobilised (µg/mg SP)
0.063	0.047	73.1	13.1	60.0	60	81	18.8
0.059	0.041	69.6	10.3	59.3			
0.068	0.055	77.8	17.4	60.4			

SEM micrographs

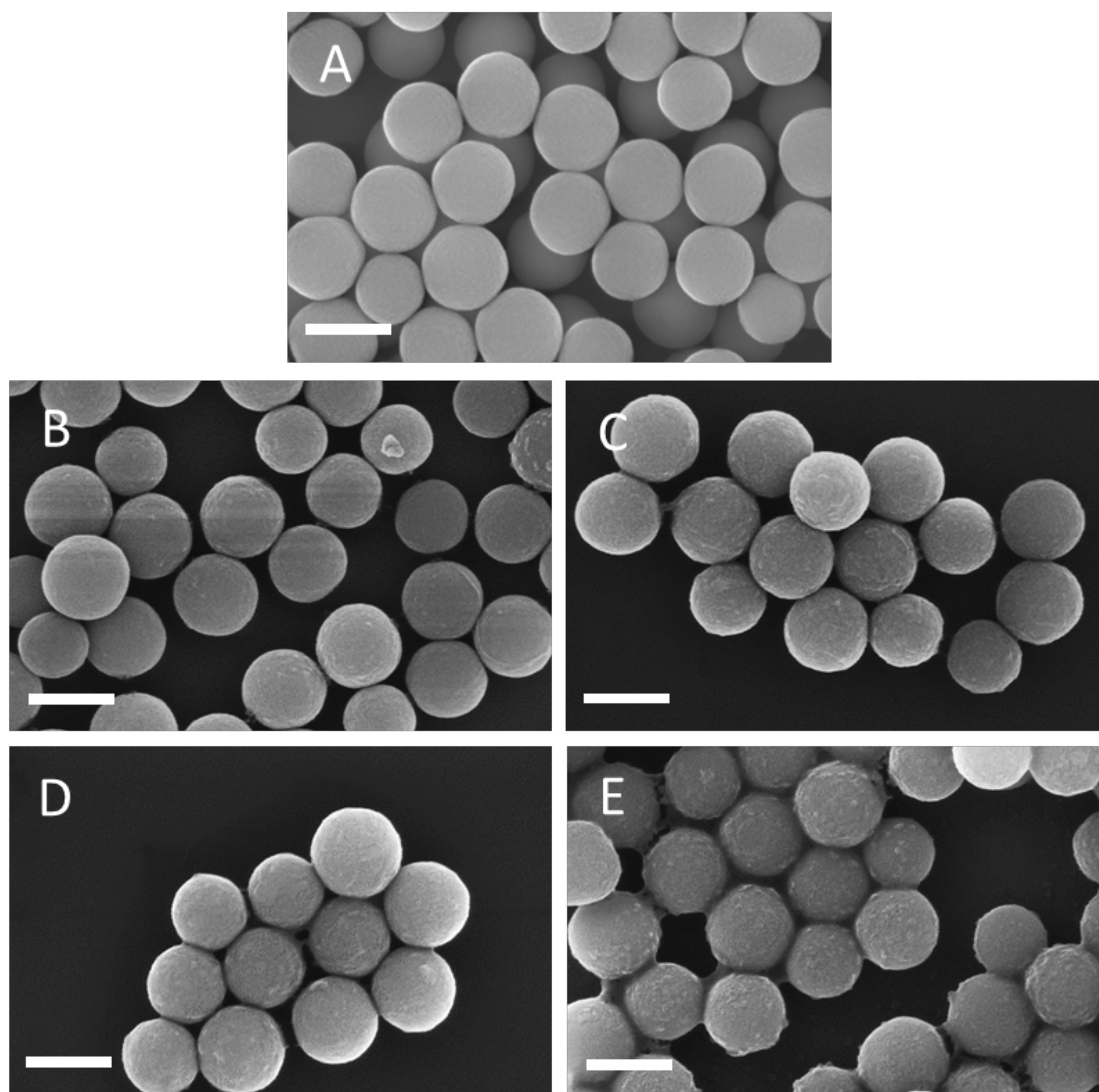

Fig. S6 SEM micrographs of bare SPs (A) and SP- Lip_{MRD9}-OS_{REF} after (B) 30 min, (C) 60 min, (D) 90 min and (E) 120 min of layer growth reaction yielding 4.0, 8.1, 10 and 13.3 nm layer thickness, respectively. Each thickness value represents the average of at least 100 SPs measured. Scale bars represent 300 nm.

Layer growth kinetics

Table S3 Layer growth kinetics of SP- Lip_{MRD9}-OS_{CD} results summary

Reaction time (min)	0	30	60	90	120
Mean diameter (nm)	290 ± 20	297.8 ± 19.7	305.1 ± 20.5	310.2 ± 19.9	315.8 ± 19.5
Layer thickness (nm)	-	3.9	7.5	10.1	12.9

Table S4 Layer growth kinetics of SP- Lip_{MRD9}-OS_{REF} results summary

Reaction time (min)	0	30	60	90	120
Mean diameter (nm)	290 ± 20	297.9 ± 20.5	306.3 ± 19.4	310.1 ± 19.9	316.7 ± 19.7
Layer thickness (nm)	-	4.0	8.1	10	13.3

References

1 T. Klemetsen, *et al.*, *Nucleic Acids Res.*, 2018, **46**, 692-699
2 S. Roda, *et al.*, *Angew. Chem. Int. Ed.*, 2022, **61**, e202207344